DOCUMENT RESUME

ED 373 065 TM 021 785

AUTHOR	Nelson, David E.; Lawrence, Barbara J.
TITLE	Utah's Major Student Assessment Programs.
INSTITUTION	Utah State Office of Education, Salt Lake City.
PUB DATE	Mar 94
NOTE	34p.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Achievement; *Achievement Tests; *Criterion Referenced Tests; *Educational Assessment; Educational Policy; Elementary Secondary Education; Item Banks; National Competency Tests; Policy Formation; Standardized Tests; *State Programs; *Testing Programs; Test Results; Test Use; Writing Tests
IDENTIFIERS	Educational Indicators; *Performance Based Evaluation; *Utah

ABSTRACT

The purpose and results of the state testing programs in Utah are described. At present there are seven major assessment programs conducted under the leadership of the Utah State Office of Education. These include the three components of the Core Assessment Program: Criterion-Referenced Tests and item pools, Writing Assessment, and the new Performance Assessment Modules. These components focus heavily on helping teachers identify the specific core curriculum concepts that have and have not been mastered by individual students, and they provide information on strengths and weaknesses of the curriculum and instruction. The Statewide Testing Program, the Utah Educational Quality Indicators effort, and the State-by-State National Assessment are all programs that focus on assessing the quality of the total system in providing relevant information to the public and policymakers on levels of academic achievement. A final program, District Performance Reports, allows school districts to profile their unique characteristics and academic performance for the benefit of their public, policymakers, and staff. An exhibit and three charts present testing results. (Contains 4 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TM

ED 373 065

Utah's Major Student Assessment Programs

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

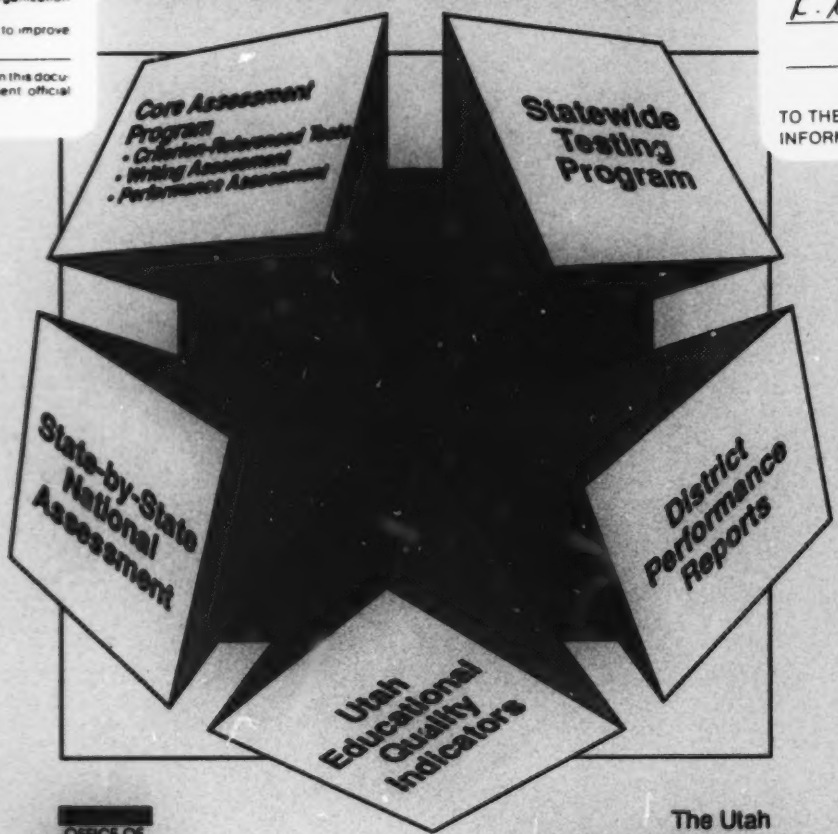
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. RAPHAEL

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."



The Utah
State Office of Education

250 East 500 South
Salt Lake City, Utah
84111

Scott W. Bean
State Superintendent
of Public Instruction

BEST COPY AVAILABLE

TM Ca 11785

UTAH'S MAJOR STUDENT ASSESSMENT PROGRAMS

Utah State Office of Education

Scott W. Bean
State Superintendent of Public Instruction

Jerry P. Peterson
Associate Superintendent, Instructional Services

David E. Nelson
Director, Evaluation and Assessment

by

David E. Nelson and Barbara J. Lawrence

March 1994

UTAH STATE BOARD OF EDUCATION
UTAH STATE BOARD FOR APPLIED TECHNOLOGY EDUCATION

District 1
Neola Brown
Beaver, Utah 84713

District 2
Lynn Haslem
Altonah, Utah 84002

District 3
Harold S. Jensen
Gunnis[illegible] Utah 84634

District 4
Keith T. Checketts
Logan, Utah 84321

District 5
Thomas F. Davidson
Ogden, Utah 84403

District 6
Milton Kendrick
Ogden, Utah 84403

District 7
Katharine B. Garff
Bountiful, Utah 84010

District 8
Daryl C. Barrett
Salt Lake City, Utah 84108

District 9
Kay McDonough
Magna, Utah 84044

District 10
Donald G. Christensen
West Valley City, Utah 84120

District 11
C. Grant Hurst
Sandy, Utah 84093

District 12
Boyd F. Jensen
Murray, Utah 84107

District 13
Allen E. Litster
Midvale, Utah 84047

District 14
Marlon O. Snow
Orem, Utah 84057

District 15
Linnea S. Barney
Orem, Utah 84058

Scott W. Bean, Executive Officer

Twila B. Affleck, Secretary

FOREWORD

Over the last two decades, the State Board and State Office of Education have put into place a variety of student assessment programs for purposes ranging from assessing concept mastery of individual students to gauging the academic effectiveness of Utah's entire system of public education. The purpose of this brochure is to acquaint educators and policymakers with the full scope of student assessment programs currently operational at the state level.

At the present time, there are seven major assessment programs which are conducted under the leadership of the State Office of Education. These include the three components of the Core Assessment Program; Criterion-Referenced Tests and Item Pools, Writing Assessment, and the new Performance Assessment modules. The focus of the three Core Assessment Program components is heavily on helping teachers identify the specific core curriculum concepts which have and have not been mastered by individual students as well as providing information on strengths and weaknesses of curriculum and instruction in specific areas. The Statewide Testing Program, the Utah Educational Quality Indicators effort, and the State-by-State National Assessment are all programs which focus on assessing the quality of the total system in providing relevant information to the public and policymakers on levels of academic achievement. A final program, District Performance Reports, allows school districts to profile their unique characteristics and academic performance for the benefit of their public, policymakers, and staff.

We strongly believe that Utah's comprehensive and multi-dimensional approaches to student assessment represent a central ingredient in providing a high quality education to each Utah student.

Scott W. Bean
State Superintendent
of Public Instruction

UTAH'S MAJOR STUDENT ASSESSMENT PROGRAMS

TABLE OF CONTENTS

Page #

Core Curriculum Assessment Programs

I. Core Criterion-Referenced Tests and Item Pools 1

II. Core Writing Assessment 7

III. Core Performance Assessment Modules 9

Total System Assessment Programs

IV. The Utah Statewide Testing Program 13

V. The Utah Educational Quality Indicators Program: "The State's Report Card" 17

VI. State-By-State National Assessment 19

VII. District Performance Reports 21

Toward a Balanced View of Student Assessment 25

I. THE CORE CRITERION-REFERENCED TESTS AND ITEM POOLS

The Utah State Office of Education, in partnership with many other education organizations, has mounted an ambitious program of development of measurement strategies and tools to support Utah school districts in assessing student performance based on Utah's Core Curriculum. With the implementation of the State Graduation Requirements of 1984, the Utah State Board of Education launched on a new path which had significant implications for both curriculum and assessment. Utah's 1984 graduation requirements put in place a statewide core curriculum for the state of Utah which set out major goals and objectives for virtually every area of the curriculum in grades kindergarten through twelve. These same graduation requirements also called upon Utah school districts to accomplish various kinds of assessment in relationship to the state core.

Utah's Core Assessment Program is probably the nation's most comprehensive effort in developing and implementing a wide range of criterion-referenced testing. One of the truly remarkable aspects of this enormous development effort is the strong collection of education organizations which has contributed to this test construction program.

Under the leadership of the State Office of Education, a program of test development and refinement was initiated in 1985 which has involved the state's colleges and universities, in-state and out-of-state research organizations and virtually every Utah school district. This test development program is ongoing. As the state core curriculum is revised, new series of tests are being developed to reflect the new core.

This cooperative effort has included a program of measurement construction which has had two major aspects. These include a program of construction of intact, end-of-

level, and end-of-course tests in mathematics, reading, and science. Supplementing this program of test construction, the Core Assessment Program has also embarked on a program of construction of pools of test items in virtually every area and level of the curriculum. Both the intact, criterion-referenced tests and the item pools are coded specifically to the Utah Core Curriculum. The following is a brief description of the two major components of the core measurement program:

1. End-of Level/End-of-Course Tests. The State Office of Education has been working with several Utah school districts, Utah colleges and universities, as well as state and national research organizations in developing end-of-level and end-of-course tests in mathematics, reading, and science. These tests are developed with great technical precision and are field tested at least three times. The end-of-level and end-of-course tests have two major purposes: first, they provide a final check on student attainment of core curriculum content; second, they help document program strengths and weaknesses.

2. Test Item Pools. The other major type of product being developed in conjunction with Utah school districts and universities, as well as other groups, is a set of test item pools for use by teachers and districts in more informal, day-to-day or unit-oriented student assessment. Development of these item pools is completed or underway in virtually every area of the curriculum and at every level. The unique aspect of these test item pools is that they are completely coded to the state core curriculum.

Core Assessment Criterion-Referenced Tests and Item Pools

· 36 elementary end-of-level, criterion-referenced tests have been developed in mathematics, reading and science. Tests exist for grades 1-6 for each subject.	· 42 secondary end-of-course, criterion-referenced tests have been developed for 21 courses in mathematics and science.	· 28 test item pools have been developed for virtually every area of the core curriculum. Each item pool contains 800 to 3,000 test items.

1993 District Implementation of Core Assessment Tools

· Nearly 600,000 State Core Assessment Tests were administered to Utah students by districts in the spring of 1993. (up from 200,000 in 1989)	· Nearly all Utah school districts administered criterion-referenced tests based on the core curriculum.	· 37 districts used both the state core tests and state scoring. (up from 22 in 1989)

Example of a Core Assessment Student Profile for Grade 4

STUDENT PROFILE

DISTRICT/AREA

TEST ELEM MATH GRADE FOUR

GRADE LEVEL 04

TEST NUMBER 00000MA0404-A

DATE OF TESTING 1993

NUMBER TESTED

SCHOOL/CLASS

STUDENT/REPORT I.D

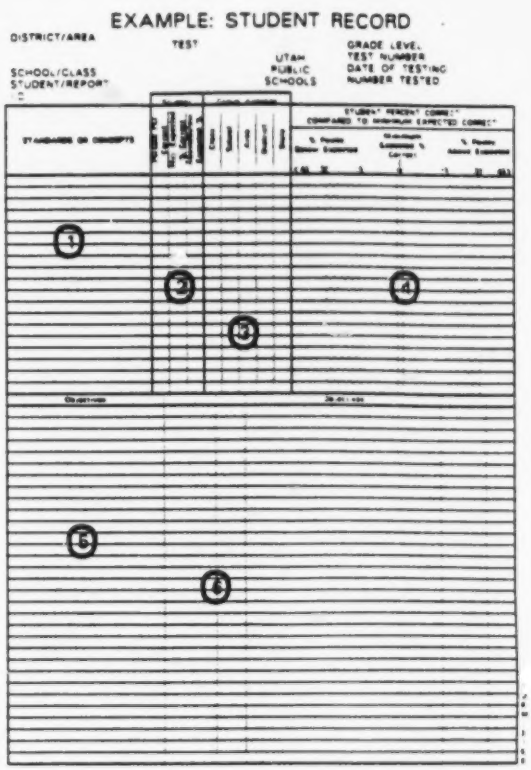
UTAH PUBLIC SCHOOLS

STANDARDS OR CONCEPTS	Student: Percent (%) Correct	Student: Minimum Expected % Correct	Student: Above/Below Expected %	Group Average: Class	Group Average: School	Group Average: Area	Group Average: District	Group Average: State	STUDENT PERCENT CORRECT COMPARED TO MINIMUM EXPECTED PERCENT CORRECT (< 45 · 30 · 15 · ↓ · 15 · 30 · 45 >)
WHOLE NUM/PLACE VALUE	100	60	+		84		84		
COMPARISON/ROUNDING	60	60	+		84		84		
ADDITION/SUBTRACTION	20	60	-		83		83		
MULTIPLICATION/FACTS	80	60	+		92		92		
MULTIPLI/REGROUPING	40	60	-		87		87		
DIVISION FACTS	100	60	+		92		92		
DIV/DIGIT DIVISOR	00	60	-		77		77		
FRACTIONS/CONCEPT OF	60	60	+		96		96		
FRACTIONS/ADD/SUBTRACT	40	60	-		85		85		
GEOMETRY	80	60	+		85		85		
MEASUR/LENGTH/AREA/TIM	40	60	-		72		72		
PROBLEM SOLVING	25	50	-		63		63		
GRAPH	100				63		63		
TOTAL TEST	55				83		83		

Objectives	[illegible]	[illegible]	Objectives	[illegible]	[illegible]
WHOLE NUM/PLACE VALUE	5/5	100			
COMPARISON/ROUNDING	3/5	60			
ADDITION/SUBTRACTION	1/5	20			
MULTIPLICATION/FACTS	4/5	80			
MULTIPLI/REGROUPING	2/5	40			
DIVISION FACTS	5/5	100			
DIV/DIGIT DIVISOR	0/5	0			
FRACTIONS/CONCEPT OF	3/5	60			
FRACTIONS/ADD/SUBTRACT	2/5	40			
GEOMETRY	4/5	80			
MEASUR/LENGTH/AREA/TIM	2/5	40			
LENGTH	2/3	67			
AREA	0/1	0			
TIME	0/1	0			
PROBLEM SOLVING	1/4	25			
GRAPH	1/1	100			

[illegible] BY UTAH STATE OFFICE OF EDUCATION

Example of a Core Assessment Student Profile for Grade 4



1 Standards or Concepts:
A description of the various sections of the test.

2 Student:

Percent (%) Correct
This calculation reflects the percentage of items correct on this specific section of the test.

Minimum Expected % Correct
This score is the minimum percent correct. It is established as a guideline for displaying competency on each section of the test. This score has been selected as a result of expert judgement and/or research on the test.

Above/Below Expected %
+ (Plus) Sign = The student's score was equal to or above the Minimum Expected % Correct.
- (Minus) Sign = The student's score was below the Minimum Expected % Correct.

3 Group Averages:

Class
The average percent correct for the teacher's class in which the student is enrolled.

School
The average percent correct for all students in the school (at a grade level or in a course) who took the test.

Area
The average percent correct for all students in an area (at a grade level or in a course) of the district who took the test.

District
The average percent correct for all students in the district (at a grade level or in a course) who took the test.

4 Student's Percent Correct Compared to Minimum Percent Correct:
A graph showing the difference between the student's percent correct and the expected percent correct.

5 Objectives:
A list of the educational objectives measured by the test.

6 Number Correct/Possible and Percent Correct:
Shows the number of items that the student answered correctly for each objective compared to the number of items possible. The student's percent correct for each objective is also indicated.

UTAH CORE ASSESSMENT
Test Item Pools

ELEMENTARY

Mathematics (1 volume)
- Grades 1-6

Reading (2 volumes)
- Primary - Grades 1-3
- Intermediate - Grades 4-6

Music (1 volume)
- Grades K-6

Social Studies (2 volumes)
- Grades 1-6
- Supplement to Grade 1

Visual Arts (1 volume)
- Grades K-6

Science (2 volumes)
- Primary - Grades 1-3
- Intermediate - Grades 4-6

Library Media (1 volume)
- Grades K-6

SECONDARY

Mathematics (5 volumes)
- Middle School and Junior High School Edition (Mathematics - Level 7, Algebra Preparation and Mathematics - Level 8, Elementary Algebra)
- Basic High School Courses Edition (Elementary Algebra, Intermediate Algebra, Geometry)
- Applied Math I and Applied Math II
- Advanced Courses Edition (Advanced Algebra, Trigonometry, Analytic Geometry, Calculus)

Social Studies (2 volumes)
- Middle Schools/Junior High Courses
- Senior High Courses

Music (3 volumes)
- Levels 7-8
- Instrumental Music, 9-12
- Musicianship and Chorus, 9-12

Visual Arts (1 volume)
- Foundations 1

Music Observation Assessment Tools (2 volumes)
- Instrumental Music
- Musicianship & Chorus

Science (7 volumes)
- Middle School and Junior High School Edition (Life Science, Earth-Space, Physical Science)
- Senior High Courses (Biology, Biological Earth, Physical Earth, Chemistry, Human Biology, Physics)

Vocational Education (3 volumes)
- Business and Marketing (Self-Assessment & SEP)
- Home Economics & Health Occupations (Self-Assessment & SEP)
- Industrial Arts & Agriculture (Self-Assessment & SEP)

II. CORE WRITING ASSESSMENT

In conjunction with the Northwest Regional Educational Laboratory, the State Office of Education has mounted a large development effort in the area of creating resources for the direct assessment of writing and other aspects of language arts. This program has included both the development of products to facilitate direct writing assessment as well as the provision of major inservice experiences for teachers and administrators.

The centerpiece of the Utah approach to direct assessment of writing is a six-trait analytic scoring model. This model supports the scoring of student writing on six dimensions. These include support products and manuals as well as inservice that is directly focused on building teacher capability to score writing on these six dimensions.

The state has sponsored several statewide writing assessment workshops to equip district-level personnel and teachers with the capability to implement the Utah Writing Assessment Model as well as skills and products which enable those personnel trained to carry on the training back at the district and school levels.

The following are the major products which have been created as part of the writing assessment program:

Large-Scale Model Writing Assessment

This is a prepackaged direct writing assessment intended for large-scale use that includes all the materials necessary to administer the assessment--student writing booklets with prompts, student directions, teacher directions, letters for schools, etc. The assessment is based on a model in which writing is described and analyzed along six dimensions or traits: ideas, organization, voice, word choice, sentence fluency, and conventions.

District Writing Assessment Handbook

For those who do not want to use the prepackaged materials described above, this handbook provides the guidance needed to design one's own large-scale direct writing assessment. Included are discussions of the various decisions to be made, as well as many sample scoring guides, letters, reports, and newspaper articles.

Integrating Writing Assessment and Instruction: A Five-Day Training Institute for Teachers

This handbook is used to provide teachers with direct experience on how to analyze student writing systematically and how to incorporate assessment into writing instruction. The focus is on empowering teachers to be better assessors of student writing and on empowering students to be self-assessors. Included are instruction on the six-trait analytical model, sample student papers, and many instructional ideas.

Prompt Bank

"Prompts" are writing topics that can be assigned to students to prompt them to write. They can be used in a large-scale assessment as well as providing a resource for instruction. The collection contains almost 200 prompts covering grades K-12 and is cross-referenced to the Utah core curriculum objectives in writing.

Writing Assessment Training Video

This video presents a one-day training on how to assess student writing using the six-trait model. The video is interactive--video-based presentation of information is interspersed with local participant discussions and scoring exercises.

Speaking and Listening Assessment Resources

The State Office of Education maintains a library of sample assessment instruments for speaking and listening. These are cross-referenced to core curriculum objectives.

Additional Writing Assessment Training

Training is available for both teachers and districts interested in doing writing assessment. Contact Mary Beth Clark at the Utah State Office of Education for further information.

III. CORE PERFORMANCE ASSESSMENT MODULES

In the effort to reform America's public schools, performance assessments have emerged as an effective tool for students, teachers, and administrators. Profiles Corporation, on behalf of the Utah State Office of Education, has developed a series of performance assessments for science, mathematics, and social studies in grade levels one through six. Performance assessment stresses the measurement of students' abilities to apply what they have learned to solve problems (such as mathematics problems or science experiments) and/or create products (such as essays or art work).

The framework presented in the Utah core curriculum provided the foundation for the new performance assessments. The assessments sample extensively from and are representative of the standards and objectives of the Utah core curriculum. These assessments will serve as a companion to the Statewide Testing Program and criterion-referenced Core Assessment Program that are already in place. Good assessment practice combines information from a number of different methods of assessment to provide a comprehensive picture of how students are performing, where group strengths and weaknesses lie, and what needs to be done to improve education for students. The addition of high-quality performance assessments to Utah's already state-of-the-art assessment program truly puts Utah on the cutting edge of the effort to reshape assessment and support instructional reform in our educational system.

In all, 90 performance assessments have been developed and extensively field tested--five in each subject area at each grade level (1-6). As with the criterion-referenced Core Assessment tests, these assessments will be made available to districts for use, at their discretion, to enhance the district assessment program.

Each assessment uses an interesting, real-world scenario, and formats vary from activity to activity. The assessments take on more of the flavor of a classroom activity than that of a test. In addition, depending on how the school/district intends to use the results, teachers are free to structure the administration of the assessments to best meet their needs and the needs of the students. For example, administration of the assessment may be spread across several days, students may work alone or in groups, and teachers can provide clarification or additional information to students as needed to assist them in completing the task. This flexibility in administration of the performance assessments may greatly reduce the anxiety students often experience when taking traditional multiple choice tests.

Each performance assessment has three components:

- The **Student Assessment**, which is designed so that students will be able to write and draw in the assessment booklet itself, thus making it self-contained with no loose parts to misplace.

- **Teacher Directions**, which include a list of the content area, concepts/skills, processes, and critical thinking skills sampled in the assessment. Also included is a description of the scenario, a scripted dialogue for generating discussion, and step-by-step directions for administering the exercise.

- A **Scoring Guide**, which provides an item-by-item scoring rubric for evaluators. For each item, the guide describes and gives examples of answers appropriate to each point level for the item.

Materials needed for the activities will be items any classroom is stocked with, such as colored pencils or crayons, glue, and scissors. Scoring guides are clear and easy to use, which allows teachers to be an integral part of the evaluation process without making excessive demands on their time or the school system's funds.

As with the Core Assessment criterion-referenced tests, camera-ready copies of all performance exercises have been made available to each district and school in the state. It will then be the responsibility of the districts and/or schools to reproduce copies for teachers' use. The one-sided originals and a coding system will make it easy for teachers to prepare to use the assessments in their classes and to maintain a master file.

A training video is being produced to provide guidance and information necessary to help teachers use the assessments effectively. In addition to instructions regarding administration and scoring of the assessments, the video includes segments showing the actual administration of the exercises during the field test. The video will be distributed along with the camera-ready copies of the assessments.

A variety of orientations and training workshops are being planned to help implement this performance-based approach to assessment of students.

Core Performance Assessment Program

Curriculum Area	Modules Ready for Use in:											
	1993-94						1994-95					
	Grade						Grade					
	1	2	3	4	5	6	1	2	3	4	5	6
Mathematics	5	5	5	5	5	5						
Science	5	5	5	5	5	5						
Social Studies	5	5	5	5	5	5						
Art							2	2	2	2	2	2
Reading							5	5	5	5	5	5

Five Modules for Each Subject for Each Grade Level (2 in Art)

BEST COPY AVAILABLE

IV. THE UTAH STATEWIDE TESTING PROGRAM

Background

1993 is the fourth year of implementation for Utah's Statewide Testing Program. The 1990 Utah Legislature passed a set of bills dealing with testing and accountability. Specifically, this legislation mandated the administration of the same norm-referenced achievement test statewide to students in grades five, eight, and eleven. Public reporting of school and district averages from this program was also specified in these laws.

Formal responsibility for creating and planning this program was vested with the State Board and State Office of Education. The State Superintendent, in conjunction with school district superintendents, appointed the State Testing Committee which accomplishes detailed planning and establishes guidelines for conducting the Statewide Testing Program. The State Testing Committee is made up of representatives of ten of Utah's forty school districts as well as staff from the State Office of Education.

The fourth administration of tests under this legislative mandate occurred in the fall of 1993. In most schools, testing took place from the last week in September through the first two weeks of October. The test chosen by the State Testing Committee for administration in this program was the Stanford Achievement Test. Through a legislative appropriation, all tests, answer sheets, and other testing materials are purchased by the State Office of Education and provided to local school districts and schools for administration. Likewise, the State Office of Education assumes all costs of scoring and reporting for the tests administered at grades five, eight, and eleven.

Reporting of Results

Initial results, such as individual student profiles and school profiles, are made available to school districts by mid-November each year. These are the same kinds of reports which have typically been produced through the test scoring service at the State Office of Education. In December, a series of special reports are produced and sent to each local school district for them to share with their board of education and public. These reports feature both median scores obtained by schools and districts as well as predicted score ranges which allow a comparison between a school median and a range of scores which would be expected in schools with similar students. These reports use this same report format to present results for each district and school.

State Results for Major Subtests

The attached table presents the statewide results based on the administration of the Stanford Achievement Test to Utah fifth, eighth, and eleventh graders in the fall of 1990, 1991, 1992, and 1993. Results are presented for all major subtests including mathematics, reading, language/English, science, social science, and the total basic battery. While the Stanford has many additional specific subscores, these major subtests are a good indicator of performance in the five major curriculum areas measured by the test as well as the total basic battery.

Utah Statewide Testing Program - Fall 1993

STATE RESULTS FOR MAJOR SUBTESTS

Median National Percentile Ranks for the Total State

Stanford Achievement Test, Eighth Edition

Subtest	Grade 5				Grade 8				Grade 11			
	1990	1991	1992	1993	1990	1991	1992	1993	1990	1991	1992	1993
Mathematics	60th	62nd	62nd	60th	53rd	54th	55th	53rd	54th	59th	59th	59th
Reading	53rd	55th	53rd	53rd	55th	55th	55th	55th	58th	58th	61st	58th
Language/English	48th	48th	48th	48th	45th	45th	45th	45th	45th	51st	51st	51st
Science	52nd	56th	56th	56th	53rd	53rd	58th	58th	60th	60th	60th	60th
Social Science	55th	55th	55th	55th	50th	50th	54th	54th	56th	56th	56th	56th
Total Basic Battery	53rd	55th	54th	54th	51st	51st	53rd	51st	53rd	55th	56th	56th
Students Tested	34,870				36,411				30,135			

The National Norm is 50 for each subtest

V. THE UTAH EDUCATIONAL QUALITY INDICATORS PROGRAM: "THE STATE'S REPORT CARD"

Since 1967, the Utah State Office of Education has been preparing this report series. The Utah Educational Quality Indicators Report Series has brought together valuable information on student performance and attitudes from a variety of sources including the American College Testing Program, the Scholastic Aptitude Test, the Advanced Placement Program, the Statewide Testing Program, the census, and unique sources of information available at the time of publication of each individual report.

The Quality Indicators Series provides a valuable supplement to the Utah Statewide Testing Program in its emphasis on examining performance measures which are collected independent of the educational system. The focus of the program is on the examination of trends in performance over extended periods of time. The program features timely single source reports as well as extensive and in-depth analyses across several measurement programs.

The Quality Indicators Series
Individual Reports and Multiple Sources Documents
o American College Testing Program (ACT) Information o Scholastic Aptitude Test (SAT) Results o Preliminary Scholastic Aptitude Test (PSAT) Results o Utah Statewide Testing Program Results o Advanced Placement Program Results o Adult Educational Attainment in Utah o Courses Taken by Utah Senior High School Students o National Assessment and International Evaluation Study Findings

VI. STATE-BY-STATE NATIONAL ASSESSMENT

The National Assessment of Educational Progress (NAEP) has existed since 1969. NAEP was established by the Congress to collect and report information on what American students know and what they can do. It is the nation's only very long-term, comparable, and representative assessment of student achievement. NAEP test items are written around a framework prepared for each content area assessed which represents the consensus of groups of curriculum experts, educators, and members of the general public.

In the late 1980s, Congress authorized National Assessment to include states and territories to participate voluntarily in a series of trial state-by-state assessments testing a sufficient number of students in each state to produce accurate information for that state. In 1992, the State of Utah participated in the national State-by-State Assessment of mathematics at grades four and eight and reading at grade four. Some 44 states and territories took part in the 1992 State-by-State National Assessment of both mathematics and reading. The assessment is based on the selection of stratified random samples of schools and students. Utah had 193 schools in the assessment across both grade levels. Utah's overall student participation rate of approximately 95 percent was among the highest in the nation. In total, 2,799 fourth graders and 2,726 eighth grade students were tested in the state.

Results from the 1992 State-by-State Assessment of Mathematics showed that Utah students at both grade four and grade eight scored substantially better than their peers nationally. Utah fourth grade students had an average overall mathematics proficiency score of 223. This contrasted with the national figure of 217 and ranked the

state 12th among the 44 states and territories participating. Utah eighth graders demonstrated an average proficiency of 274 compared to the national figure of 266. At this grade level, the state was 8th in the nation.

An analysis of results for all of the 44 states and territories which participated in the 1992 study showed that students in only two states at the fourth grade and three states at the eighth grade scored statistically significantly higher than Utah students. Utah's performance also exceeded the western regional average and the performance of other participating western states.

A unique feature of National Assessment reporting is the identification of specific achievement levels used to communicate levels of mathematics proficiency. These three levels include basic, proficient, and advanced. In the 1992 mathematics study, 67 percent of Utah fourth graders and 72 percent of Utah eighth graders scored at or above the basic level. The basic level represents competent performance in mathematics knowledge and skills.

Twenty percent of Utah fourth graders and 27 percent of the state's eighth graders scored at or above the proficient level. Both the percentage of Utah students at the basic level and the percentage at the proficient level were substantially higher than percentages of students tested across the nation.

National Assessment plans to continue its program of state-by-state trial assessments with an assessment of reading at grades four, eight, and twelve in 1994.

VII. DISTRICT PERFORMANCE REPORTS

House Bill 170, passed by the 1990 Utah Legislature, requires that Utah school districts submit an annual performance report to the State Board of Education and the Legislature by January 15 of each year. Copies of the district performance reports are to be distributed to the residence of each student enrolled in every district.

In addition to the action of the Utah legislature, there is broad-based support for the implementation of school district performance reports in the State of Utah. The State Board of Education has directed Utah State Office of Education staff to, "Provide models, training, and guidelines to enable all Utah school districts to prepare and distribute to the public, legislature, and the State Board of Education, an annual planning and performance report."

The initiative by the State Board is also complimented in formal statements by the Utah School Superintendent's Association and the Governor. The Utah School Superintendent's Association has endorsed a position paper which, among other things, states that Utah school districts are committed to student assessment programs which include: "a reporting program which provides individual student achievement test information to parents, school test results to the school's community, and district-wide test data to the general public." The Governor has indicated that reporting to parents and to the public should be a critical activity to be accomplished by school districts.

Many Utah school districts have over several years compiled and reported district-wide performance information in a systematic way. This ranged from some extremely sophisticated approaches which included careful selection of information to be reported, targeting of specific audiences and attractive formats to very basic or highly technical

elaborations of various performance measures. Virtually all school districts currently have access to several different relevant kinds of performance and descriptive information. Thus, in most instances the major challenge is the systematic organization analysis and reporting of available information.

It is not the intent of this program to limit in any way what school districts feel is important to report to their public. Districts are encouraged to communicate information which is timely and relevant whether or not such information is specified by the law. The focus of the handbook published by the Utah State Office of Education to facilitate the creation of reports is on identifying some key statistical information which should have relevance to the public, policymakers, and educators. The handbook does address all required information specified in H.B. 170. Creative approaches to presenting narrative and other statistical information are encouraged.

Data Elements Specified By H.B. 170 For Inclusion in District Performance Reports

I. **Achievement** - Trend Data
- A. District Norm-Referenced Test Scores
- B. ACT Scores / Percent Taking Test
- C. A.P. Scores / Number of Students Taking Tests / Number Passing
- D. Criterion-Referenced Test Scores

II. **Students**
- A. Enrollment Trends
- B. Ethnic Distribution of Student Population
- C. Attendance Trends
  - 1. A.D.A. (If Available)
  - 2. A.D.M.
- D. Drop Out Rates and Trends

III. **Fiscal Information**
- A. Revenue By Source and Total
- B. Expenditures by Category and Total
- C. Current Expenditures Per Student and Trends
  - 1. District
  - 2. State
  - 3. National

IV. **Staff**
- A. By Level (Number and Percent). Level = Elementary, Junior High, Senior High, and District Office.
- B. Pupil/Teacher Ratios and Trends
- C. Average Years of Experience of Professional Staff

V. **Curriculum**
- A. Course-Taking Patterns (Trends Based on State Data)

VI. **School Buildings**
- A. Number of Buildings by Level
- B. Number of Year-Round Schools
- C. Number of Extended Day Schools

VII. **Assistance to Students and Families**
- A. Percentage of Free School Lunch Participants
- B. Percentage of Students Who Receive Fee Waivers
- C. Total Dollar Amount of Fee Waivers
- D. Percentage of Students Whose Families Receive AFDC Funds

VIII. **Support Services** - DISTRICT DISCRETION
- A. School Lunch
- B. Student Transportation
- C. Maintenance
- D. Security

TOWARD A BALANCED VIEW OF STUDENT ASSESSMENT

Not since the late 1960s has the world of educational measurement seen the debate which is currently taking place over methodology in measuring educational achievement. At the crux of the debate is the ongoing dialogue concerning the relevance and appropriateness of three major classes of student assessment techniques. These include norm-referenced tests, criterion-referenced tests, and performance assessments which are also termed, "authentic" assessment in some quarters.

Understanding the present discussion concerning these three approaches to measuring educational achievement demands that we know something of the origin of each. Norm-referenced tests first appeared in the early 1920s with the publication of the first edition of the Stanford Achievement Test.[1] Interestingly, norm-referenced tests were a reaction to the common educational assessment practices of the day which consisted principally of student essays and oral examinations conducted by teachers. Subsequently, a very large industry has grown up around the development and scoring of norm-referenced tests, as has a very sophisticated technology for the construction and interpretation of such tests.

In 1963, Robert Glaser, writing in American Psychologist, coined the term, "criterion-referenced test."[2] Whereas norm-referenced tests rely on relating the scores of individuals or groups to a national norm group, criterion-referenced tests focus on measuring the specific objectives or concepts which students have or have not attained. Both norm-referenced tests (NRTs) and criterion-referenced tests (CRTs) are typically composed of multiple choice test items.

What many believe to be the newest entry on the stage of achievement testing is typically called performance assessment. Performance assessment measures are actually much older than either norm-referenced or criterion-referenced tests (essay examinations are a kind of performance assessment). The focus in performance assessment is on having students create products either through the written word or, as is the case of the arts, a non-verbal product such as a painting.

Profiling the Characteristics of Alternative Achievement Measures

Exhibit 1 presents a comparison of norm-referenced tests, criterion-referenced tests, and performance assessment in terms of six different comparison points. The comparison points include what each approach uses as a **reference point**, **specificity**, **quality of test**, **types of scales**, **curriculum relevance**, and **sensitivity to instruction**.

In respect to a **reference point**, we have noted that norm-referenced tests are based on comparisons with a national norm group of students in the same grade. Criterion-referenced tests are linked to specific objectives or concepts and this is often done in a quantitative fashion. The performance assessment may be linked to either quantitative or qualitative performance standards. For example, a student product such as a rendering in the third dimension may be compared to standards representing different levels of proficiency in interpreting the third dimension.

In the area of **specificity**, norm-referenced tests tend to be survey measures and are, therefore, a global picture of the performance of students or schools. Criterion-referenced tests, on the other hand, tend to be more specific and identify attainment or non-attainment of rather specific curriculum objectives. Performance assessments are often very specific. An example would be student skill in writing a persuasive essay given

Exhibit 1

A COMPARISON OF NORM-REFERENCED TESTS, CRITERION-REFERENCED TESTS, AND PERFORMANCE ASSESSMENTS

Comparison Point	Criterion-Referenced Tests (CRTs)	Norm-Referenced Tests (NRTs)	Performance Assessments
Reference Point	Performance Standard	Students in Same Grade	Quantitative and/or Qualitative Performance Standard
Specificity	More Specific	More Global	Very Specific to Global
Quality of Test	Variable	Typically High Quality	Extremely Variable
Types of Scales	Few	Many	Few/None
Curriculum Relevance	High	Variable	Extremely High
Sensitivity to Instruction	High	Moderate	Extremely High

rather explicit instructions. This is not always true, however, since vehicles such as student portfolios which are rather global are also often characterized as being in the realm of performance assessment.

Whereas norm-referenced tests are almost always carefully constructed and designed to meet very exacting **psychometric standards**, the same is not true of all criterion-referenced tests or performance assessments. Criterion-referenced tests may indeed have the same high standards of measurement quality exhibited by norm-referenced tests, but this is not necessarily the case. When criterion-referenced tests are to be used for important purposes, it is critical that the user looks carefully at evidence of the test's quality. The same caution would apply to performance assessments. Outside the domain of writing assessment, little has been done in demonstrating reliability and validity of performance assessments. Thus, with both CRTs and performance assessments, the operative phrase is, "caveat emptor."

In the area of **scaling**, most of us are familiar with the many scales used with norm-referenced tests. These include metrics such as national percentile ranks, standard scores, grade equivalents, etc. Criterion-referenced tests are almost always interpreted as a percentage correct for individual students or an average percentage correct for groups. Performance assessments are often reported on some kind of descriptive scale where, hopefully, the numerical scale points have defined meanings. This is another area of performance assessment which needs considerable work.

The **curriculum relevance** of norm-referenced tests is variable. In areas such as mathematics and reading there is often excellent relevance to a given local or state curriculum. This is not necessarily true for score areas such as science or social studies. Criterion-referenced tests, by definition, have high curriculum relevance since they are

built to measure the objectives of a specific local, state, or national curriculum. Performance assessments take curriculum relevance to its highest point. Often there is little or no difference between actual curriculum and the exercise being used as a performance assessment. A good example would be a writing exercise where the student is asked to write a brief persuasive essay. The difference between curriculum and assessment here is not apparent.

Related to curriculum relevance is the concept of **sensitivity to instruction**. Since norm-referenced tests are a survey of what students are expected to know, they can only be expected to be moderately sensitive to instruction. Criterion-referenced tests, on the other hand, should show high sensitivity to instruction since they are based on the same objectives that define the curriculum. Performance assessments, again, score at the highest level on this dimension since they are most like instruction itself.

An Argument for a Balanced Approach to Measuring Educational Achievement -- Avoiding the Holy War

An analysis of the characteristics of the three alternative approaches to measurement of educational achievement suggests that each of these measures has legitimacy in its own right. Norm-referenced tests constitute an excellent vehicle for accomplishing high quality measurement across very broad areas of several content domains and doing so in a fashion which is economical in the use of time as well as quite inexpensive in actual dollar outlay. The norm-referenced test gives us a snapshot of student performance or system performance and lets us relate that picture to a national standard of performance--the national norm. Thus, the NRT has meaning to parents and policymakers in a way which may never be achieved by criterion-referenced testing or performance assessment.

Well-constructed criterion-referenced tests such as those built for the Utah Core Curriculum Assessment Program as end-of-level or end-of-course tests, also have a definite role to play in measuring student achievement. CRTs target specific areas of student strength and weakness and provide highly diagnostic information which lets us address the needs of individual students as well as groups of students. Strong curriculum relevance and sensitivity to instruction make multiple choice criterion-referenced tests ideal as guides for instruction. Such tests also tend to be inexpensive in terms of student time as well as purchase and scoring costs.

The use of performance assessment techniques lends a reality to the measurement of educational achievement which is very difficult to attain with conventional, multiple choice procedures. While conventional NRTs and CRTs provide economy of assessment and allow large numbers of objectives to be assessed in relatively short periods of time, the strengths of performance assessments lie in their comprehensiveness and ability to emulate what might be expected in "real world" applications of what students know. Unfortunately, they do so at relatively high cost in time and scoring. In a performance assessment environment, the assessment becomes very much a part of instruction. Thus, it is no surprise that performance assessment scores so high on dimensions such as curriculum relevance and sensitivity to instruction.

From all of the preceding, it seems clear that any search for the one optimal method of achievement assessment which is equally valid and useable for all assessment situations is pointless. Indeed, those who promote any one of the three major options which have been discussed in this paper as the only answer to educational assessment are likely rendering a great disservice to the profession. Writing in the December 1991 issue of Phi Delta Kappan, Maeroff has the following observation concerning performance

assessment as a panacea for all assessment applications:[3]

> There is an inclination among proponents to regard alternative (performance) assessment as suitable for all purposes--diagnosis, selection, and accountability at all levels. It may be that an alternative assessment that is a marvelous indicator of an individual child's academic progress will prove fairly useless for other purposes. Americans may have to decide whether comparisons are what they seek in alternative assessment or whether they prefer to use the approach for other, more individualized purposes. (Page 276)

A more productive approach to measurement of educational achievement would find educators understanding the strengths, weaknesses, and appropriate applications of each of the three major approaches to assessment of achievement. The following illustrations represent some of the appropriate applications for each of the three techniques:

- **Norm-referenced tests**. The NRT represents a cost-effective, technically sophisticated means of surveying achievement information of both individuals and large groups of students. NRTs provide broad overviews of student performance which are valid general indicators of either individual or group status. Such tests also represent a rather economical approach which enables education to keep parents, the public, and policymakers informed concerning some key attributes of school achievement.

- **Criterion-referenced tests**. While the NRT provides the broad overview, CRTs provide economical information which allows educators to target student strengths and weaknesses especially well in the skills areas. With their close link to curriculum, CRTs at both the individual student and group levels promote in-depth understanding of achievement status which may be linked specifically to ongoing instructional decisions about programs for individuals and groups.

- **Performance assessment**. The inclusion of performance assessments in the basket of achievement assessment techniques adds a dimension of vividness and reality to what is assessed. Through performance assessment, educators can examine what students can actually create, whether those products are verbal or nonverbal. Performance assessment adds a dimension of richness to student evaluation which is obtainable in no other way.

Training in Classroom Assessment - The Missing Ingredient

What needs to happen to effect the level of understanding which currently characterizes the discussion of measuring educational achievement? It is absolutely clear that the area of measurement of student achievement has been short-changed in both preservice and inservice educational programs for teachers and administrators.[4] It seems vital that both teacher preparation institutions and those responsible for inservice training in public education take a very serious look at this problem and begin building programs to enhance the practice of achievement measurement in the schools. A hopeful note along these lines is a classroom assessment training package developed by Richard Stiggins at the Northwest Regional Educational Laboratory in Portland, Oregon. This media-based set of workshops provides an excellent beginning along a path that has hardly been traveled at all.

References

1. Ann Anastasi, Psychological Testing, Macmillan, 1976.

2. Robert Glaser, "Instructional Technology and the Measurement of Learning Outcomes," American Psychologist, 1963 18, p. 519.

3. Gene Maeroff, "Assessing Alternative Assessment," Phi Delta Kappan, December 1991, p. 270.

4. Richard Stiggins, "Assessment Literacy," Phi Delta Kappan, March 1991, p. 535.